

Random walks in nonuniform environments with local dynamic interactions

Christopher M. Baker, Barry D. Hughes,* and Kerry A. Landman

Department of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia

(Received 3 June 2013; revised manuscript received 6 August 2013; published 7 October 2013)

We consider a class of lattice random walk models in which the random walker is initially confined to a finite connected set of allowed sites but has the opportunity to enlarge this set by colliding with its boundaries, each such collision having a given probability of breaking through. The model is motivated by an analogy to cell motility in tissue, where motile cells have the ability to remodel extracellular matrix, but is presented here as a generic model for stochastic erosion. For the one-dimensional case, we report some exact analytic results, some mean-field type analytic approximate results and simulations. We compute exactly the mean and variance of the time taken to enlarge the interval from a single site to a given size. The problem of determining the statistics of the interval length and the walker's position at a given time is more difficult and we report several interesting observations from simulations. Our simulations include the case in which the initial interval length is random and the case in which the initial state of the lattice is a random mixture of allowed and forbidden sites, with the walker placed at random on an allowed site. To illustrate the extension of these ideas to higher-dimensional systems, we consider the erosion of the simple cubic lattice commencing from a single site and report simulations of measures of cluster size and shape and the mean-square displacement of the walker.

DOI: [10.1103/PhysRevE.88.042113](https://doi.org/10.1103/PhysRevE.88.042113)

PACS number(s): 05.40.Fb, 87.10.Mn

I. INTRODUCTION

The study of random walks can be traced back to a letter to Nature in 1905 by Karl Pearson [1]. Since then, many variations of the problem have been considered, both with a variety of applications in mind and also for their mathematical interest [2–5]. Although random walk models have found applications in the physical sciences in such disparate contexts as polymer physics [6], quantum computing [7], and xerography [8], Pearson's motivating example was biological (mosquito dispersal) [9]. Biology continues to provide a rich source of opportunities for random walk modeling, from describing the migratory and settling behavior of invertebrates [10] to probing the migration of individual cells [11]. The variant of the random walk model discussed in this paper, involving walks in restricted domains, was inspired by biology. In tissue, cell motility is impeded by a support structure called the extracellular matrix (ECM) [12]. Some cells are able to produce enzymes called matrix metalloproteinases, which are used to break down the ECM and allow greater cellular movement [13–15]. Interpreting a motile cell as a lattice-based random walker and regions of intact matrix as blocked sites produces the simple generic random walk problem that we pose below and describe as the *erosion model*. A realistic model of the cell-ECM interaction would necessarily be somewhat more complex, but our simple erosion model will be seen to have some interesting properties.

In the erosion model, we assign a state to each lattice site. A site is either *allowed* or *blocked*. At each time step, the walker attempts one move, either left or right with equal probability. If the move is onto an allowed site, then it is always successful. However, if the step is onto a blocked site, the walker must change the state of the site from blocked to allowed if it is to move there. The walker is successful in unblocking the site with probability p_s (the *snipping probability*). The state of the

site is then permanently changed to allowed. If the walker is unsuccessful in clearing the target site, then the attempt to step onto the target site is unsuccessful. The walker remains in its original position and the site remains blocked. An example of a walker attempting to move right is illustrated in Fig. 1.

The erosion model can be implemented on any lattice, but except for Sec. V our discussion here is restricted to the one-dimensional case (the linear chain). The mean-square displacement for an unbiased one-dimensional random walk in an unbounded environment is

$$\langle R_n^2 \rangle = 2Dn, \quad (1)$$

where D is the diffusivity [3]. For the classical Pólya random walk (nearest neighbor stepping only, with the two possible steps each having probability $1/2$), we have $2D = 1$, while if the walker has a probability $p_m/2$ of stepping left, a probability $p_m/2$ of stepping right, and a probability $1 - p_m$ of pausing at each time step, we have $2D = p_m$. One question we investigate here is whether the result (1) also holds for erosion models, with a p_s -dependent diffusivity. A second question of interest is the growth of the interval length (that is, the number of sites in the cluster of allowed sites). We seek information about

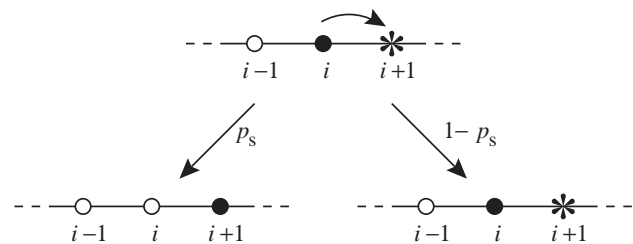


FIG. 1. The site occupied by the walker is shown in black and the star denotes a blocked site. A walker at position i attempts to move right onto the blocked site $i + 1$. It is successful with the probability p_s , in which case the walker makes the step, and $i + 1$ becomes an allowed site. With probability $1 - p_s$ the walker fails to clear this site $i + 1$, and the move is aborted, leaving the walker at site i .

*barrydh@unimelb.edu.au

the random variable L_n , which we define as the length of the interval after n steps, or the random variable T_l , which we define as the time taken for the interval size to be l for the first time.

We always commence the process by placing a walker on an allowed site, but we consider several initial conditions, starting with the simplest, in which the walker starts at an isolated allowed site, with all other sites of the lattice blocked (Sec. II). We show how the mean and variance of the random variable T_l can be calculated exactly and demonstrate that in this system fluctuations are never negligible. The standard deviation is asymptotically proportional to the mean as $l \rightarrow \infty$. Although we are unable to determine exact formulas for the mean length $\langle L_n \rangle$ of the allowed interval and the mean-square walker displacement $\langle R_n^2 \rangle$ at a given time, simulation reveals the asymptotic behavior of both of these, and we develop a mean-field type approximation for $\langle L_n \rangle$.

In Secs. III and IV we start the process at an occupied site in a realization of a random lattice of allowed and blocked sites. In both sections, the initial interval size (which is random) is determined by declaring sites at random to be allowed with probability p and blocked with probability $1 - p$, so that we have a realization of the Bernoulli site percolation process in one dimension [4,16,17]. (Conventionally, in percolation theory one would describe as “occupied” what we call “allowed” sites, but that terminology would create ambiguities in the discussion.) Beyond one dimension, statistical attributes of interest in percolation theory as a function of p , such as the mean cluster size, the distribution of cluster sizes, and the probability that a given site lies on an infinite cluster, are extremely hard to calculate, though they have been extensively studied [4]. In contrast in one dimension, all of the calculations are straightforward and short, and we now discuss briefly several useful results.

The probability that a randomly chosen site is part of a cluster of precisely m allowed sites, with $m > 0$, is given by $mp^m(1-p)^2$ [4]. This follows since we require m contiguous sites being allowed (probability p^m), a blocked site on either end [probability $(1-p)^2$], and have m possible places within the cluster at which the chosen site may be located. In our model, the walker always starts on a cluster at the origin, so the origin is always an allowed site. Computing the conditional probability that the origin is on a cluster of size m , given that it is on a cluster of size at least one, we have our required distribution of the initial cluster size C_p :

$$\Pr\{C_p = m\} = mp^{m-1}(1-p)^2 \quad (m \geq 1). \quad (2)$$

Within this cluster, each site has probability $1/m$ of being the starting site. The mean initial cluster size is

$$\langle C_p \rangle = \sum_{m=1}^{\infty} m \times mp^{m-1}(1-p)^2 = \frac{1+p}{1-p}. \quad (3)$$

Here and in a number of results derived below, finite and infinite series are evaluated directly by recognizing them as differentiated geometric progressions and we do this without further comment.

Although the same initial size distribution (2) for the cluster on which the walk commences is used in Secs. III and IV, we treat the remainder of the lattice differently in these two

sections. In Sec. III all sites that are not part of the cluster of allowed sites on which the walk commences are set as blocked, so that the random walker has to erode never-ending blocked regions and the length of the allowed interval can only ever increase by one on any successful erosion step. We are able to adapt the methods of Sec. II to compute the mean time to grow the allowed cluster to a given length, but again we are only able to compute $\langle R_n^2 \rangle$ by simulation.

When p_s is set to zero, the model of Sec. III reduces to a random walk on a restricted domain (the finite interval between the two nearest blocked sites on either side of the starting site) with reflecting boundary conditions [2]. Previous work on this problem has focused on the probability distribution of the position of the walker and first passage times [18,19]. In our case we consider a slightly different problem, where the size of the region is chosen randomly in each realization and depends on the parameter p . We also seek the mean-square displacement, rather than the probability distribution and first passage time to a blocked site. Some exact and empirical results for the case $p_s = 0$ are given in the Appendix.

In Sec. IV we consider the most challenging one-dimensional problem, in which the walker starts at a random allowed site in a random lattice with allowed site density p . In this case, as well as the cluster of allowed sites on which the walker starts, there are additional clusters of allowed sites separated from it by intervals of blocked sites. Some successful erosion steps will remove a single blocked site that separates two clusters of allowed sites and so join the clusters together. Consequently, the length of the allowed cluster on which the walker moves may have a large increase on a single time step.

The basic erosion idea can be implanted on an arbitrary lattice. We have emphasized the one-dimensional case, where a number of exact or approximate analytic calculation can be performed and compared against simulations, but simulations are easily performed on other standard lattices. As an illustration, in Sec. V we have considered the three-dimensional simple cubic lattice, confining our attention to the case where the walker commences from a single allowed site.

Techniques used in the paper include exact probability arguments, approximate or heuristic probability arguments, and simulations. Random variables are always capitalized and averages are denoted by angle brackets. Simulations reported were performed using MATLAB or MATHEMATICA, except for the diffusivity estimates in Sec. IV, where C++ was used (with the default random number generator replaced by the Mersenne Twister [20]).

While there has been a considerable body of work on random walk or diffusion processes in static random environments [21], there appears to have been little previous work on random walk problems for which one or more walkers modify the geometry of their environment as they move. In 1983, Herrmann [22] posed and simulated the problem of the “moles’ labyrinth”, which in our terminology would correspond to seeding an infinite lattice at random with a given density of simultaneous random walkers. All sites not occupied by a walker at the seeding time are deemed disallowed. The walkers are given probability 1 of converting any disallowed site onto which they step to allowed status, and Herrmann’s interest was in the phase transition from disconnected finite clusters of

allowed sites to a state in which an infinite connected cluster appears.

II. ONE-DIMENSIONAL EROSION FROM A SINGLE SITE

A. Time to grow a cluster to a given size

We consider the special case of an erosion problem in which the walker starts on a single allowed site, where all other sites of the one-dimensional lattice are initially deemed blocked. The walker undertakes a simple, unbiased random walk, where an attempt to step onto a blocked site succeeds with probability p_s (thereby changing that site's status to allowed), while if the attempt fails, the walker remains at its previous location at the edge of the cluster of allowed sites (Fig. 1). Suppose that the size of the cluster has just increased from $l - 1$ to l . Let I_k denote the indicator random variable for the event that it takes precisely k attempts to visit either or both of the blocked sites at 0 and $l + 1$ to convert one of these sites to an allowed site (and so increase the size of the cluster of allowed sites from l to $l + 1$). Then

$$\langle I_k \rangle = \Pr(I_k = 1) = p_s(1 - p_s)^{k-1}.$$

Each unsuccessful attempt to step onto a blocked site leaves the walker at the allowed end-point site from which the attempt was made. Hence, the times between attempts are independently, identically distributed random variables. We denote their mean by τ_l and their variance by σ_l^2 . Therefore, the time $T_{l,k}$ taken for k attempts has mean $\langle T_{l,k} \rangle = k\tau_l$ and variance $\text{Var}(T_{l,k}) = k\sigma_l^2$. The random time taken to increase the length of the cluster from l to $l + 1$ is

$$T_l^+ = \sum_{k=1}^{\infty} I_k T_{k,l}.$$

Although this is a formal infinite series, in each realization of the random process, only a single term is nonzero. The calculation of the mean time taken to increase the cluster size is straightforward:

$$\langle T_l^+ \rangle = \left\langle \sum_{k=1}^{\infty} I_k T_{k,l} \right\rangle = \sum_{k=1}^{\infty} p_s(1 - p_s)^{k-1} k\tau_l = \frac{\tau_l}{p_s}. \quad (4)$$

To compute the variance of T_l^+ we note first that since $\langle I_k I_m \rangle = 0$ if $k \neq m$ and $\langle I_k^2 \rangle = \langle I_k \rangle$, we have

$$\begin{aligned} \langle (T_l^+)^2 \rangle &= \left\langle \left(\sum_{k=1}^{\infty} I_k T_{k,l} \right) \left(\sum_{m=1}^{\infty} I_m T_{m,l} \right) \right\rangle \\ &= \sum_{k=1}^{\infty} \langle I_k \rangle \langle (T_{k,l})^2 \rangle \\ &= \sum_{k=1}^{\infty} p_s(1 - p_s)^{k-1} [\text{Var}(T_{l,k}) + \langle T_{l,k} \rangle^2] \\ &= \sum_{k=1}^{\infty} p_s(1 - p_s)^{k-1} [k\sigma_l^2 + k^2\tau_l^2] \\ &= \frac{\sigma_l^2}{p_s} + \frac{(2 - p_s)\tau_l^2}{p_s^2}. \end{aligned} \quad (5)$$

Hence, we find that

$$\text{Var}(T_l^+) = \langle (T_l^+)^2 \rangle - \langle T_l^+ \rangle^2 = \frac{1}{p_s^2} [p_s\sigma_l^2 + (1 - p_s)\tau_l^2]. \quad (6)$$

It is not difficult to calculate τ_l and σ_l^2 using simple difference equation arguments (cf. [3], pp. 114–115). For example, $\tau_l = t(1) = t(l)$, where $t(k) = k(l + 1 - k)$ is the unique solution of the difference equation $t(k) = 1 + [t(k - 1) + t(k + 1)]/2$ with boundary conditions $t(0) = t(l + 1) = 0$. We have

$$\tau_l = l \quad \text{and} \quad \sigma_l^2 = \frac{1}{3}l(l^2 - 1). \quad (7)$$

Consequently,

$$\langle T_l^+ \rangle = \frac{l}{p_s} \quad \text{and} \quad \text{Var}(T_l^+) = \frac{1}{3p_s^2} [p_sl^3 + 3(1 - p_s)l^2 - p_sl].$$

If we now define T_l to be the time taken for a cluster of size 1 to grow to a cluster of size l we have

$$\langle T_l \rangle = \sum_{k=1}^{l-1} \langle T_k^+ \rangle \quad \text{and} \quad \text{Var}(T_l) = \sum_{k=1}^{l-1} \text{Var}(T_k^+).$$

Using the identities $\sum_{k=1}^{l-1} k = \frac{l(l-1)}{2}$,

$$\sum_{k=1}^{l-1} k^2 = \frac{l(l-1)(2l-1)}{6}, \quad \sum_{k=1}^{l-1} k^3 = \frac{l^2(l-1)^2}{4},$$

we are able to conclude that

$$\langle T_l \rangle = \frac{l(l-1)}{2p_s} \quad (8)$$

and

$$\text{Var}(T_l) = \frac{l(l-1)}{12p_s^2} [p_sl^2 + (4 - 5p_s)l - 2]. \quad (9)$$

As anticipated, the time to grow to a given interval size l goes to infinity as the snipping probability $p_s \rightarrow 0$.

In the random walk literature, the number of distinct sites visited by a walk (sometimes called the *range* of the walk) has been a problem of considerable interest [3]. In one dimension, the length of the interval that has been visited is known as the *span* of the walk. In the special case $p_s = 1$, our results give us the mean and variance of the time taken to reach a given span. The determination of the probability distribution for the time T_l is more difficult. Even for the special case $p_s = 1$ this appears to be an unsolved problem. We have not pursued this here.

B. Cluster size and mean-square displacement at a given time

Instead of considering the time T_l to a given allowed interval size l , one might like to know the allowed interval size L_n after n steps. Since we have $2p_s\langle T_l \rangle \sim l^2$, we anticipate that we should have, at least as $n \rightarrow \infty$,

$$\langle L_n \rangle^2 \sim C(p_s)n. \quad (10)$$

Since Eqs. (8) and (9) show that fluctuations in T_l are comparable to the mean for all l , it is a little naive to hope that $C(p_s) = p_s$. Simulations of sufficient accuracy to clarify

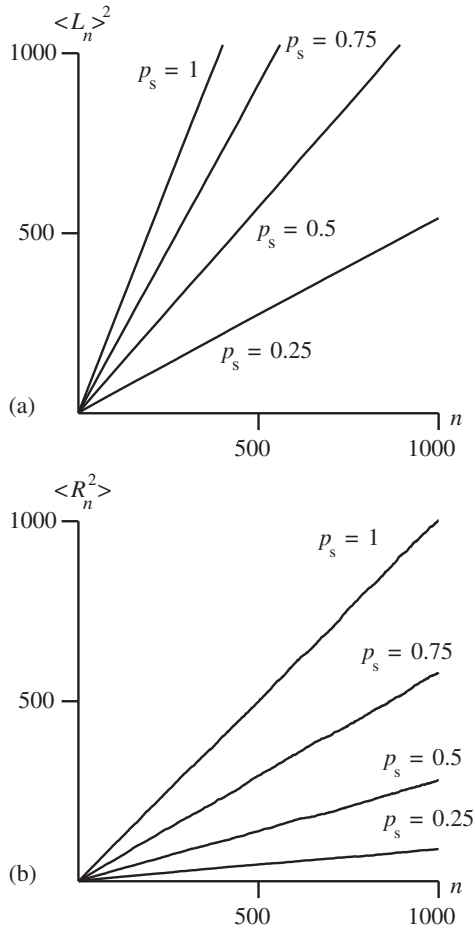


FIG. 2. The erosion model for a walker with all sites other than the starting site initially blocked for four values of the snipping probability p_s . Results shown are averaged over 10^4 simulations of 10^3 step walks. The case $p_s = 1$ corresponds to the ordinary random walk. (a) Mean length of the interval. For $p_s = 1$ our simulations are indistinguishable at this level of resolution from a straight line of slope $8/\pi$ (the asymptotic form for the square of the mean span of an ordinary random walk). (b) Mean-square displacement. For $p_s = 1$ we recover the elementary exact result $\langle R_n^2 \rangle = n$.

this point are easily performed. In Fig. 2(a) we show the time evolution of the mean allowed interval length $\langle L_n \rangle$. For each value of p_s , we generated 10^4 walks of 10^3 steps with MATHEMATICA. The excellence of the conjectured asymptotic form (10) is clear. Moreover, for $p_s = 1$, the simulation data are indistinguishable from the relation $\langle L_n \rangle^2 = 8n/\pi$, which is the (large- n asymptotic) exact result for the span of a random walk [5] and equivalently, the asymptotic form of $\langle S_n \rangle^2$, where S_n is the number of distinct sites visited [3].

In Fig. 2(b) we show the time evolution of the mean-square displacement $\langle R_n^2 \rangle$. It appears that as $n \rightarrow \infty$

$$\langle R_n^2 \rangle \sim 2D(p_s)n; \quad (11)$$

that is, the walker has a well-defined diffusion constant. From the simulation we observe that $2D(1) = 1$, correctly recovering the familiar result for an unconstrained simple random walk.

In Fig. 3 we show estimates of the coefficients $C(p_s)$ and $2D(p_s)$ computed from the final length and the final

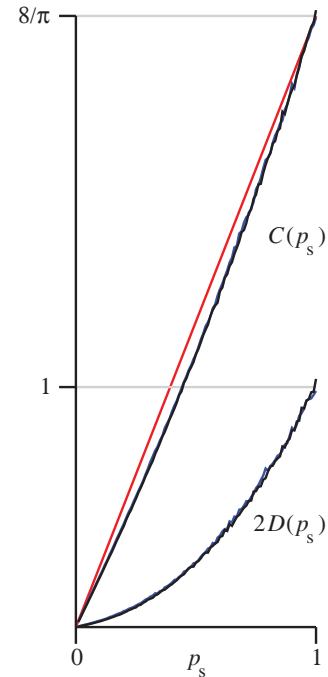


FIG. 3. (Color online) The erosion model for a walker with all sites other than the starting site initially blocked. For $p_s = 0.01, 0.02, \dots, 1$ we show estimates of the coefficients $C(p_s)$ and $2D(p_s)$. The estimates are obtained by assuming that Eqs. (10) and (11) are equalities at the largest value of n used. The results for 10^3 step walks, shown as blue (gray) curves, and for 10^4 step walks, shown as black curves, are virtually indistinguishable. In each case 10^4 individual walks were averaged. The gray horizontal lines indicate the exact value for $p_s = 1$. The approximation $C(p_s) \approx 8p_s/\pi$ is shown as a red (gray) line.

mean-square displacement in 10^3 step and 10^4 step walks (10^4 realizations for each value of p_s , increments in p_s of 0.01). As the estimates based on walks of two different durations coincide closely, it appears unnecessary to consider longer walks. The coefficient $C(p_s)$ is approximated roughly by $C(p_s) \approx 8p_s/\pi$ and fits quite well, as shown in Fig. 3.

We have used the MATHEMATICA function NonlinearModelFit to fit the simulation data for $2D(p_s)$ from 10^4 realizations of 10^4 step walks with a quadratic polynomial that respects the requirements that $D(0) = 0$ and $D(1)$. We find that

$$2D(p_s) \approx ap_s + (1-a)p_s^2, \quad a \approx 0.068824 \quad (12)$$

(with R^2 goodness of fit value 0.999584). We have not plotted this empirical fit in Fig. 3, because it matches the data too closely to be distinguished from it. Despite the quality of the fit, it is important not to overinterpret global phenomenological relationships. One can obtain slightly better global fits by choosing a cubic constrained to take the correct values at $p_s = 0$ and 1, or less physically natural functional forms as $p_s/(p_s - a \ln p_s)$ or $a \exp(bp_s + c \ln p_s)$, where a , b , and c are best-fit coefficients.

The success of the simple approximation (12) suggests that we try to improve on the rough approximation $C(p_s) \approx 8p_s/\pi$ by trying the most general quadratics and cubics conforming to the exact values $C(0) = 0$ and $C(1) = 8/\pi$. MATHEMATICA

gives as the best fit of this kind (with corresponding R^2 goodness of fit values)

$$C(p_s) \approx 1.99565 p_s + 0.550825 p_s^2 \quad (R^2 = 0.99973); \quad (13)$$

$$C(p_s) \approx 2.01961 p_s + 0.478944 p_s^2 + 0.0479208 p_s^3 \quad (R^2 = 0.99974). \quad (14)$$

These empirical fits to the data raise the possibility that $C(p_s) \sim 2p_s$ as $p_s \rightarrow 0$. We demonstrate this analytically below, but we do not at present have an argument to predict the coefficient of p_s^2 in the small- p_s expansion of $C(p_s)$.

C. Mean-field theory of cluster size

Since the cluster size grows by at most one per time step, the evolution of the cluster size is governed by the exact probability relation

$$\begin{aligned} \Pr\{L_{n+1} = l\} &= \Pr\{L_n = l\} \Pr\{L_{n+1} = l \mid L_n = l\} \\ &+ \Pr\{L_n = l-1\} \Pr\{L_{n+1} = l \mid L_n = l-1\}. \end{aligned} \quad (15)$$

In the Appendix we discuss the time evolution of the position of a walker in a fixed interval (corresponding to a starting interval of nonzero length, with $p_s = 0$) and note that there is exponentially fast relaxation to a uniform distribution. Hence, especially for $p_s \ll 1$, it is natural to take as a first approximation that at any instant in the present problem, the walker's position is uniformly distributed, and thus for $l \geq 2$,

$$\Pr\{L_{n+1} = l \mid L_n = l-1\} \approx \frac{2}{l-1} \times \frac{1}{2} \times p_s. \quad (16)$$

The first factor on the right is an approximation for the probability that the walker is at the cluster edge. The other factors account exactly for the probability that the walker attempts to step beyond the cluster and that this attempt succeeds, growing the cluster length by 1. Using this approximation, Eq. (15) becomes (for $l \geq 2$)

$$\begin{aligned} \Pr\{L_{n+1} = l\} &\approx \left(1 - \frac{p_s}{l}\right) \Pr\{L_n = l\} \\ &+ \frac{p_s}{l-1} \Pr\{L_n = l-1\}. \end{aligned} \quad (17)$$

If we multiply Eq. (17) by $l-1$ and sum from $l = 2$ to ∞ , we obtain without further approximation

$$\langle L_{n+1} \rangle \approx \langle L_n \rangle + p_s \langle L_n^{-1} \rangle. \quad (18)$$

We know from Jensen's inequality that $\langle L_n \rangle \langle L_n^{-1} \rangle \geq 1$, with equality only when L_n has a unique value. Consequently, if we make the crude approximation that $\langle L_n^{-1} \rangle \approx \langle L_n \rangle^{-1}$, and also the asymptotic approximation $\langle L_{n+1} \rangle - \langle L_n \rangle \approx (d/dn) \langle L_n \rangle$, we arrive at

$$\langle L_n \rangle \frac{d}{dn} \langle L_n \rangle \approx p_s.$$

From this we conclude that as $n \rightarrow \infty$,

$$\langle L_n \rangle^2 \approx 2p_s n, \quad (19)$$

in excellent agreement for small p_s with our observations (13) and (14). Recalling Jensen's inequality, we would expect

Eq. (19), which asserts that $C(p_s) = 2$, to underestimate the growth of $\langle L_n \rangle$. Therefore, a safer prediction would be that $C(p_s) \geq 2$, again consistent with our numerical evidence, which gave $C(p_s) > 2$ except perhaps when $p_s \ll 1$.

As we now show, one may calculate the large- n asymptotic form of $\langle L_n \rangle$ implied by the approximate evolution equation (17) exactly using generating functions. It will be convenient to write

$$\varphi_n(l) = \frac{1}{l} \Pr\{L_n = l\}, \quad (20)$$

so that

$$\begin{aligned} \sum_{n=0}^{\infty} \langle L_n \rangle \xi^n &= \sum_{n=0}^{\infty} \sum_{l=1}^{\infty} l^2 \varphi_n(l) \xi^n \\ &= \sum_{n=0}^{\infty} \sum_{l=1}^{\infty} l(l-1) \varphi_n(l) \xi^n + \sum_{n=0}^{\infty} \sum_{l=1}^{\infty} \Pr\{L_n = l\} \xi^n \\ &= \frac{\partial^2}{\partial z^2} \Phi(z, \xi) \Big|_{z=1} + \frac{1}{1-\xi}, \end{aligned} \quad (21)$$

where

$$\Phi(z, \xi) = \sum_{n=0}^{\infty} \sum_{l=1}^{\infty} \varphi_n(l) \xi^n z^l. \quad (22)$$

Using the constraints that $\varphi_n(l) = 0$ for $l > n+1$ and that $0 \leq \varphi_n(l) < 1/l$, it is easy to show that for $0 < \xi < 1$ and $0 < z < \xi^{-1}$ we have

$$1 \leq \Phi(z, \xi) \leq \xi^{-1} (1-\xi)^{-1} \ln[(1-\xi z)^{-1}].$$

Consequently $\Phi(z, \xi)$ is singularity-free at $z = 1$ and in the neighborhood of $z = 1$.

Now Eq. (17) can be rewritten as

$$l \varphi_{n+1}(l) = (l-p_s) \varphi_n(l) + p_s \varphi_n(l-1), \quad (23)$$

where we have suppressed explicit reference to this as an approximate evolution equation [the only approximation in use is (16)]. If we multiply by ξ^{n+1} , sum from $n = 0$ to ∞ , and recall that $\varphi_0(l) = \delta_{l,1}$, we obtain

$$\begin{aligned} \sum_{n=0}^{\infty} l \varphi_n(l) \xi^n - \delta_{l,1} &= \xi \sum_{n=0}^{\infty} l \varphi_n(l) \xi^n - \xi p_s \sum_{n=0}^{\infty} \varphi_n(l) \xi^n \\ &+ \xi p_s \sum_{n=0}^{\infty} \varphi_n(l-1) \xi^n. \end{aligned}$$

Next we multiply by z^{l-1} and sum over l , finding that

$$\frac{\partial}{\partial z} \Phi(z, \xi) + \frac{p_s \xi (z^{-1} - 1)}{1-\xi} \Phi(z, \xi) = \frac{1}{1-\xi}. \quad (24)$$

Differentiating this equation with respect to z , setting $z = 1$ and using Eq. (21), we find a simple relation between the generating function for $\langle L_n \rangle$ and $\Phi(1, \xi)$:

$$\sum_{n=0}^{\infty} \langle L_n \rangle \xi^n = \frac{1}{1-\xi} \{1 + p_s \xi \Phi(1, \xi)\}. \quad (25)$$

The first-order ordinary differential equation (24) can be solved with the appropriate initial condition $\Phi(0, \xi) = 0$ using an

integrating factor, giving, after a little algebra,

$$\Phi(1, \xi) = \frac{1}{1 - \xi} \int_0^1 \exp[\lambda Q(t)] dt, \quad (26)$$

where $\lambda = p_s \xi / (1 - \xi)$ and $Q(t) = 1 - t + \ln(t)$. To find the large- n behavior of $\langle L_n \rangle$ we need only to extract the dominant asymptotic behavior of $\Phi(1, \xi)$ as $\xi \rightarrow 1^-$, corresponding to $\lambda \rightarrow \infty$. Noting that $Q(t)$ attains its maximum on the interval at $t = 1$, in the neighborhood of which we have $Q(t) = 2^{-1}(t - 1)^2 + \dots$, the usual method of Laplace argument [23] gives the rigorous conclusion that

$$\Phi(1, \xi) \sim \frac{\pi^{1/2}}{(1 - \xi)(2\lambda)^{1/2}} = \frac{\sqrt{\pi}}{(2p_s)^{1/2}(1 - \xi)^{1/2}}. \quad (27)$$

Hence, as $\xi \rightarrow 1^-$ we find that

$$\sum_{n=0}^{\infty} \langle L_n \rangle \xi^n \sim \frac{(\pi p_s)^{1/2}}{2^{1/2}(1 - \xi)^{3/2}}. \quad (28)$$

Since $\langle L_n \rangle$ is monotonic in n , we may now conclude from a Tauberian theorem [24] that as $n \rightarrow \infty$

$$\langle L_n \rangle \sim \frac{(\pi p_s)^{1/2}}{2^{1/2}} \times \frac{n^{1/2}}{\Gamma(3/2)} = (2p_s n)^{1/2}. \quad (29)$$

Thus, our rough argument leading to the approximation (19) actually identified correctly the dominant behavior of $\langle L_n \rangle$ implied by making the sole approximation (16): $\langle L_n \rangle^2 \sim 2p_s n$. An asymptotic expansion for $\langle L_n \rangle$ can be developed by introducing the new integration variable $\tau = -Q(t)$ in Eq. (26) using Watson's lemma [23] to produce an expansion of $\Phi(1, \xi)$ in powers of $(1 - \xi)^{1/2}$ and then using Darboux's theorem [23] to extract the large- n behavior of $\langle L_n \rangle$ to any desired order. We refrain from exhibiting the algebra here, but we find that subject to the approximation (16),

$$\langle L_n \rangle = (2p_s n)^{1/2} + \frac{1}{3} + o(1) \quad \text{as } n \rightarrow \infty \quad (30)$$

and correspondingly $\langle L_n \rangle^2 = 2p_s n + (2/3)(2p_s n)^{1/2} + \dots$.

III. ONE-DIMENSIONAL EROSION FROM A RANDOM INTERVAL

We now fix a value of p and start the walk at a randomly chosen site of a random interval of allowed sites, where the probability that the interval has length m is given by Eq. (2). All other sites of the lattice are blocked initially. A number of results for the limiting case in which $p_s = 0$ (no erosion) can be found in the Appendix.

A. Time to grow a cluster to a given size

We let T_l denote the first time at which the size of the cluster of allowed sites on which the walker moves contains at least l sites. This time is 0 if the cluster is initially of size l or greater, and so we have

$$\langle T_l \rangle = \sum_{m=1}^{l-1} m p^{m-1} (1-p)^2 \left[\frac{1}{m} \sum_{k=1}^m \langle T_{k,m} \rangle + \sum_{j=m+1}^{l-1} \langle T_j^+ \rangle \right], \quad (31)$$

where $\langle T_j^+ \rangle = j/p_s$ is as defined in Sec. II and the sum over k is interpreted as void if $m = l - 1$. The random variable $T_{k,m}$ is

the time taken for a walker starting at site k (with $1 \leq k \leq m$) to attempt to step beyond the boundary sufficiently often to increase the cluster size to $m + 1$. After this time, the basic approach of Sec. II applies, which is why the sum over k arises. To compute $\langle T_{k,m} \rangle$ we note from Sec. II that the expected time to the first attempt is $k(m + 1 - k)$. This attempt is successful with probability p_s and otherwise the walker is left adjacent to the boundary. Thus,

$$\langle T_{k,m} \rangle = k(m + 1 - k) + \sum_{i=1}^{\infty} i \tau_m (1 - p_s)^i p_s, \quad (32)$$

where the time $\tau_m = m$ has been discussed in Sec. II. Hence,

$$\begin{aligned} \langle T_l \rangle &= \sum_{m=1}^{l-1} p^{m-1} (1-p)^2 \sum_{k=1}^m \left[k(m + 1 - k) + \frac{m(1 - p_s)}{p_s} \right] \\ &\quad + \sum_{m=1}^{l-1} m p^{m-1} (1-p)^2 \sum_{j=m+1}^{l-1} \frac{j}{p_s}. \end{aligned} \quad (33)$$

Evaluation of the algebra gives

$$\begin{aligned} \langle T_l \rangle &= \frac{l(l-1)}{2p_s} - \frac{2p(3-p_s) + 3p^2(1-p_s) + 2p_s}{3p_s(1-p)^2} \\ &\quad + \frac{p^{l-1} Q(l, p, p_s)}{6(1-p)^2 p_s}, \end{aligned} \quad (34)$$

where

$$\begin{aligned} Q(l, p, p_s) &= 2p[3p^2(1-p_s) + 2p_s + 2p(3-p_s)] \\ &\quad + lp(1-p)[6 + 7p_s + p(12 - 13p_s)] \\ &\quad + l^2(1-p)^2[5p_s + p(6 - 8p_s)] \\ &\quad - l^3(1-p)^3 p_s. \end{aligned} \quad (35)$$

The last term on the right in Eq. (34) decays very rapidly as l increases. The other terms are more important at large l : We have the exact value of $\langle T_l \rangle$ for erosion from a single point (independent of p , and growing quadratically with l), plus an offset, independent of l , so that the mean time to grow to modest interval length is influenced by the random initial interval size.

B. Cluster size and mean-square displacement at a given time

Simulations provide data on the evolution of the mean cluster size $\langle L_n \rangle$, as illustrated in Fig. 4. We plot alongside the simulation results an empirical interpolation formula that we infer from earlier results: Since $\langle L_0 \rangle = \langle C_p \rangle = (1+p)/(1-p)$ from Eq. (2), and $\langle L_n \rangle \sim [C(p_s)n]^{1/2}$ as $n \rightarrow \infty$ from Eq. (10) for erosion from a single site, we consider an amalgam of these two as

$$\langle L_n \rangle \approx \sqrt{C(p_s)n + \frac{(1+p)^2}{(1-p)^2}}, \quad (36)$$

with $C(p_s)$ approximated by the fitted quadratic function (13). The formula (36) turns out to perform exceptionally well, as observed in Fig. 7. Therefore, the approximation (36) provides a satisfactory description of the cluster size as a function of site state probability p and snipping probability p_s .

The mean-square displacement $\langle R_n^2 \rangle$ obtained from simulation results becomes asymptotically linear in n , as given by

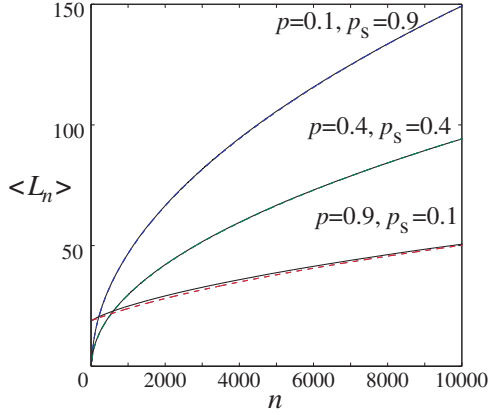


FIG. 4. (Color online) The growth over time of the cluster size $\langle L_n \rangle$ for erosion from a random interval (dashed lines) together with the empirical interpolation formula (36) (solid lines) for 10^4 steps and 2×10^4 realizations and three sets of parameter values.

$\langle R_n^2 \rangle \sim 2D(p_s)n$ [Eqs. (11) and (12)]. Therefore, as expected, the effect of the random interval size (controlled by p) fades as $n \rightarrow \infty$. The approach to linearity depends on the value of p_s . It is slowest for small snipping probabilities p_s and large values of p . This case is shown in Fig. 5.

IV. ERODING A ONE-DIMENSIONAL RANDOM ENVIRONMENT

We now consider erosion of a full realization of site percolation, with allowed sites present with probability p , starting the walk at an arbitrary allowed site. This now opens the possibility that when the walker is successful in changing a single site from blocked to allowed, two clusters of allowed sites merge to form a new allowed cluster larger than either of them. The cluster size is subject to increments of a range

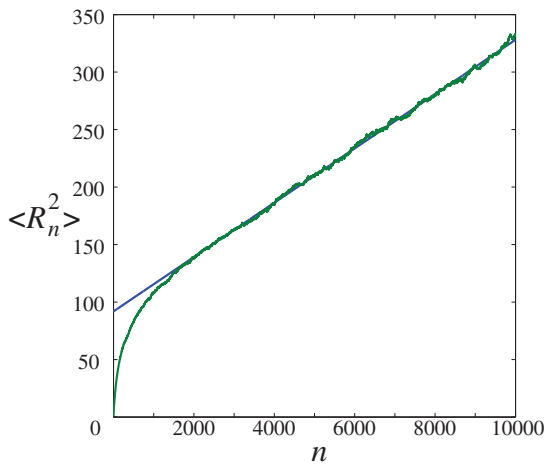


FIG. 5. (Color online) The transition to effective diffusion for erosion from a random interval. Mean-square displacement $\langle R_n^2 \rangle$ versus n for 10^4 steps and averaged over 2×10^4 realizations, with $p = 0.9$ and $p_s = 0.1$. We see that $\langle R_n^2 \rangle \sim 2D(p_s)n$ as $n \rightarrow \infty$, where the slope of the line, $2D(p_s)$, is given by the empirically determined formula (12).

of sizes. We have no exact results for this model, but we have obtained results by extensive simulation.

A. Mean-square displacement

We assume that as $n \rightarrow \infty$,

$$\langle R_n^2 \rangle \sim 2D(p, p_s)n. \quad (37)$$

To obtain acceptably accurate estimates of $2D(p, p_s)$ we generated 2×10^6 realizations of walks of 10^4 steps for each specification of p and p_s . We estimated $2D(p, p_s)$ by linear regression for $n > 2500$. To assess the likely imprecision in the estimates of $2D(p, p_s)$, we determined for the specific case $p = p_s = 0.5$ the standard error in the estimates of $\langle R_n^2 \rangle$, and performed a standard bootstrap analysis [25], which suggests that the estimate of $2D(0.5, 0.5)$ obtained has an error of $\pm 10^{-9}$.

In Fig. 6 our estimates of $2D(p, p_s)$ are illustrated. As expected, the diffusivity increases as p and p_s increase. We have considered empirical fits to the data. While many functional forms with a small number of parameters can be tuned to fit the data quite well, given the success of the empirical approximation (12) in Sec. II B the most natural fits to attempt are low-order polynomials in both variables that

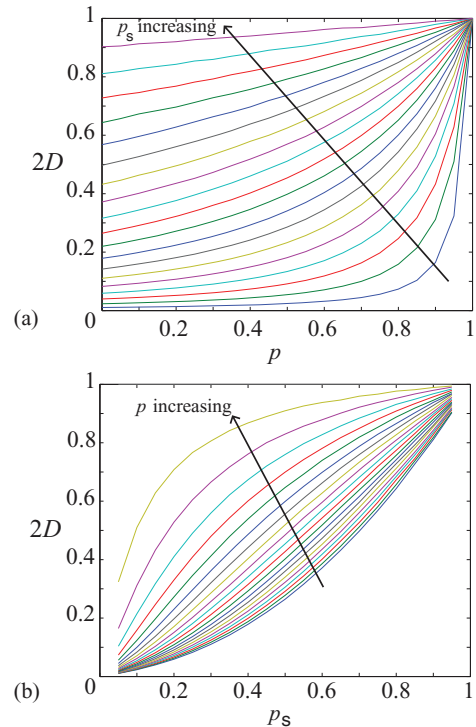


FIG. 6. (Color online) Diffusivity for eroding a random environment with initial allowed site density p and snipping probability p_s (Sec. IV) for $0.05 \leq p \leq 0.95$ and $0.05 \leq p_s \leq 0.95$ (with increments of 0.05 in each case). For each choice of p and p_s , we generated 2×10^6 random walks of 10^4 steps and the diffusivity was inferred by a linear regression fit (not constrained to pass through the origin) assuming that $\langle R_n^2 \rangle = 2D(p, p_s)n + \text{constant}$ for $n > 2500$. (a) Each curve shows $2D(p, p_s)$ for a fixed value of p_s between from 0.05 (lowest curve) to 0.95 (highest curve). (b) Each curve shows $2D(p, p_s)$ for a fixed value of p from 0.05 (lowest curve) to 0.95 (highest curve). The curves plotted are linear interpolations. Except for the lowest three curves in (a), a higher-order interpolation scheme would produce no visible difference in the figure.

respect the requirement that $2\mathcal{D}(p,0) = 0$ for $p \in [0,1)$ and $2\mathcal{D}(0,p_s) = 2\mathcal{D}(p_s)$ given in Eq. (12). These conditions give us a polynomial of the form

$$2\mathcal{D}(p,p_s) = ap_s + bpps + (1-a)p_s^2 - bpps^2, \quad (38)$$

where Eq. (12) defines the value of a as $a \approx 0.068824$. The value of b is found by fitting all the the data in Fig. 6; we find $b \approx 2.115$ with goodness of fit value $R^2 = 0.9540$. If only the $p \in [0.05,0.8]$ data in Fig. 6 are fitted with a polynomial of the form (38), then the value of b changes slightly to be $b \approx 1.716$ with an $R^2 = 0.9907$. Note that we cannot expect to find an expression for all values of p since by definition $2\mathcal{D}(1,p_s) = 1$; hence, there is a discontinuity at $p = 1$.

B. Allowed cluster size

Insight can be obtained into the difference between erosion on a random environment and erosion on a random interval by comparing the cluster size as a function of number of steps for the same parameter values p and p_s . Figure 7(a) contains such cluster size data which can be directly compared with the data in Fig. 4. The cluster size is always larger now since a successful erosion event that converts a single site to allowed status may cause two previously disjoint intervals of allowed sites to merge, so that an increase in the cluster size of more than one per erosion event can occur. We take the ratio of the

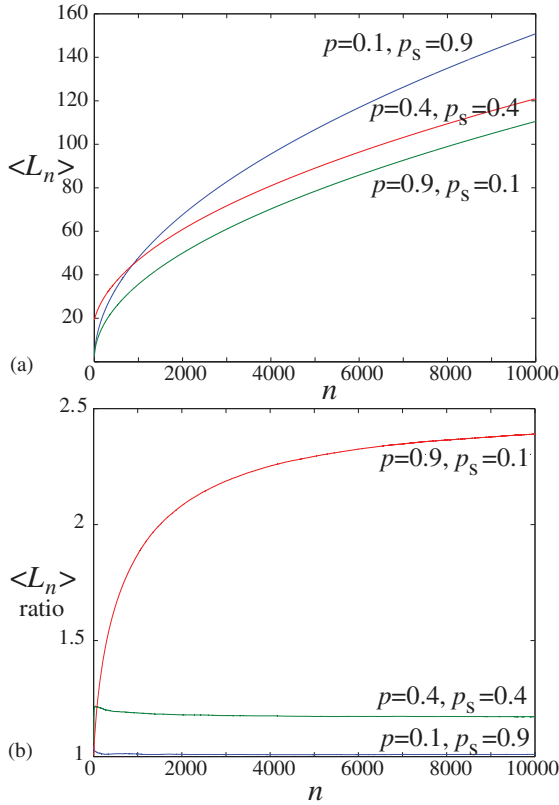


FIG. 7. (Color online) Simulation results for erosion of a one-dimensional random environment. (a) The growth over time of the mean cluster size $\langle L_n \rangle$ for 10^4 steps and 2×10^6 realizations for the same three sets of parameter values as in Fig. 4. (b) Ratio of cluster sizes in Fig. 7(a) above those in Fig. 4. The results are always greater than unity. The ratio increases as p increases.

corresponding cluster sizes and plot them in Fig. 7(b). Indeed, the average cluster size increases significantly for larger values of p and, as expected, negligibly for small values of p .

V. THREE-DIMENSIONAL EROSION FROM A SINGLE SITE

We have focused on the one-dimensional implementation of our model, where there are some exact results available, as well as useful analytic approximations. The model can be implemented on any lattice if we are prepared to forego the prospect of exact analytic results. We report some representative simulation results for the simple cubic lattice in three dimensions, which brings us closer to our original motivating context of cellular interaction with ECM.

Our simulations are all for erosion of the simple cubic lattice, starting at a single allowed site, taking $p_s = 0.25, 0.5, 0.75$, and 1 . We generated 20 000 realizations of random walks for each value of p_s considered, so that statistical fluctuations in our results are negligible at the level of resolution of the figures.

A. Growth of the allowed cluster

We have considered three measures of the state of the allowed cluster after n steps. In Fig. 8 we show the mean allowed cluster size, which we denote by $\chi(n, p_s)$. For certain erosion ($p_s = 1$), the allowed cluster size is just the number of distinct sites visited by an ordinary random walker, conventionally denoted in physics by S_n . The asymptotic behavior of the mean of S_n is well known to be [3]

$$\langle S_n \rangle \sim (1 - R)n, \quad (39)$$

where R is the probability of eventual return of the walker to the starting site. The value of $(1 - R)^{-1}$ for the standard Pólya random walk on the simple cubic lattice is known exactly in terms of the complete elliptic integral (or alternatively in terms of Γ functions) and for our purposes the numerical value $(1 - R)^{-1} \approx 1.516386$ suffices. Therefore,

$$\chi(n, 1) \sim (1 - R)n \quad \text{as } n \rightarrow \infty, \quad (40)$$

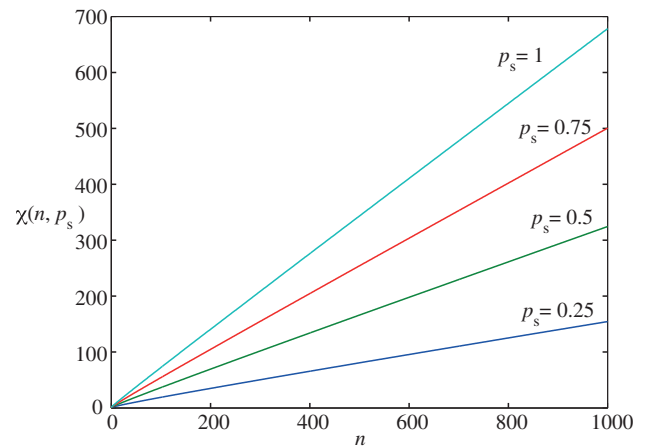


FIG. 8. (Color online) The growth of the mean cluster size $\chi(n, p_s)$ for erosion of the simple cubic lattice \mathbb{Z}^3 starting from a single site.

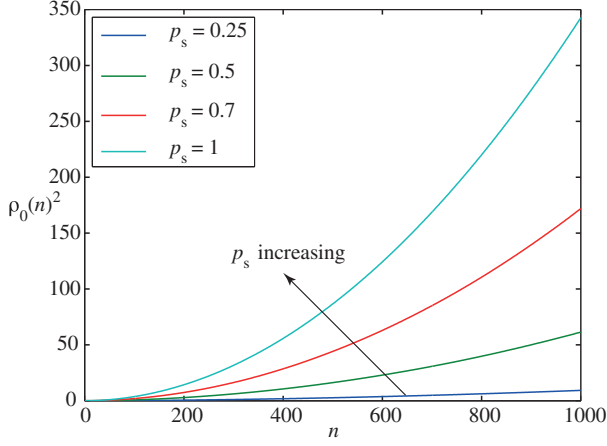


FIG. 9. (Color online) The second moment of the mass distribution about the starting position, $\rho_0(n)^2$, for the cluster of allowed sites in erosion of the simple cubic lattice \mathbb{Z}^3 starting from a single site.

where $(1 - R) \approx 0.6594627$. In our simulations, we find that $\chi(1000, 1)/1000 \approx 0.6758$, consistent with the exact result for the $n \rightarrow \infty$ limit.

For $p_s < 1$, not all attempts to move onto a disallowed site are successful, and this represents an effective random time delay on the growth of the number of distinct sites visited in a normal random walk. The transience of normal random walks in three dimensions suggests that an approximate prediction of the asymptotic form of $\chi(n, p_s)$ might be possible, but we have not yet devised anything convincing.

We have examined two measures of the geometrical size of the cluster of allowed sites at time n . We associate unit mass with each site \mathbf{r} of the cluster of allowed sites, and denote the location of the center of mass by $\bar{\mathbf{r}}$. We define

$$\rho_0(n)^2 = \left\langle \frac{\sum_{\mathbf{r}} |\mathbf{r}|^2}{\sum_{\mathbf{r}} 1} \right\rangle, \quad \rho_{\text{CM}}(n)^2 = \left\langle \frac{\sum_{\mathbf{r}} |\mathbf{r} - \bar{\mathbf{r}}|^2}{\sum_{\mathbf{r}} 1} \right\rangle, \quad (41)$$

with the sum taken over all allowed sites at time n . The positive numbers $\rho_0(n)$ and $\rho_{\text{CM}}(n)$ are the radii of gyration about the origin and about the center of mass, respectively. We show $\rho_0(n)^2$ in Fig. 9 and $\rho_{\text{CM}}(n)^2$ in Fig. 10. The well-known

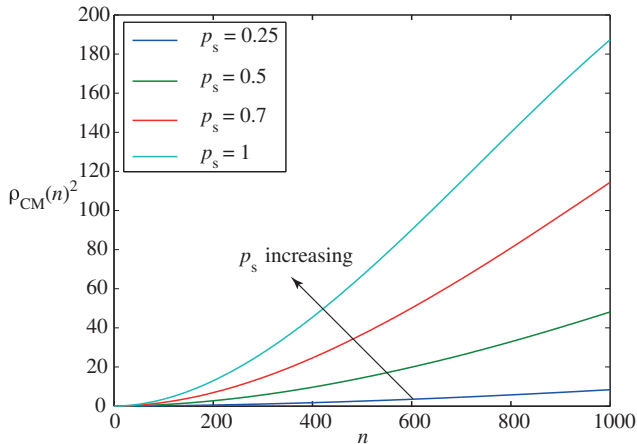


FIG. 10. (Color online) The second moment of the mass distribution about the center of mass, $\rho_{\text{CM}}(n)^2$, for the cluster of allowed sites in erosion of the simple cubic lattice \mathbb{Z}^3 starting from a single site.

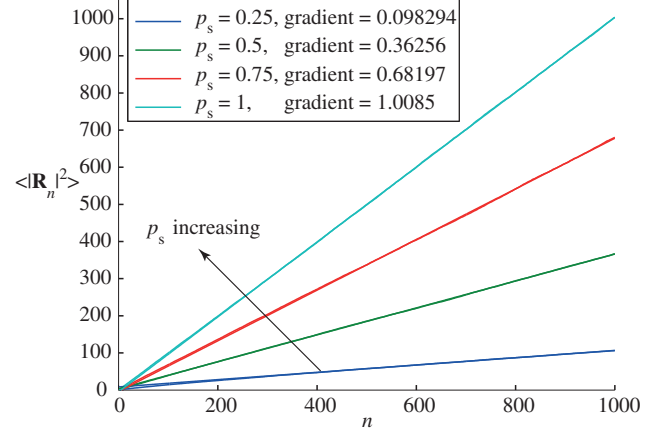


FIG. 11. (Color online) The mean-square displacement $\langle |\mathbf{R}_n|^2 \rangle$ for erosion of the simple cubic lattice \mathbb{Z}^3 starting from a single site.

asymmetry of typical realizations of ordinary random walks despite their stochastic symmetry leads to an expectation that for normal random walks $\rho_0(n)^2 > \rho_{\text{CM}}(n)^2$. Our simulations show this to be the case for the erosion model also.

B. Mean-square displacement

We conclude our brief account of the erosion of the simple cubic lattice with a discussion of the mean-square displacement of the walker. Where \mathbf{R}_n denotes the position of the walker after n steps, we show simulations of the mean-square displacement $\langle |\mathbf{R}_n|^2 \rangle$ in Fig. 11. For $p_s = 1$ we find when $n = 1000$ that $\langle |\mathbf{R}_n|^2 \rangle \approx 10008$, which compares well with the exact result that $\langle |\mathbf{R}_n|^2 \rangle = n$. The p_s dependence is captured roughly by $\langle |\mathbf{R}_n|^2 \rangle \approx p_s^\alpha n$, with $\alpha \approx 1.453$, though we not been able to predict this asymptotic form.

VI. CONCLUSION

We have considered several one-dimensional implementations of a random walk problem in which the random walker interacts with the boundaries of the interval in which the walk takes place and may enlarge the interval. For the cases in which the initial interval is a point or has a size distribution related to the cluster-size distribution in site percolation, we are able to compute exactly the mean time taken for the walker to find itself for the first time in an interval of length at least l . Although we are unable to determine the exact asymptotic growth law for the mean interval length and the mean-square displacement of the walker, using simulation and mean-field arguments we are able to identify the qualitative behavior.

For the more challenging problem in which the walker moves in a full realization of site percolation, single-site erosion events can link random intervals. We have studied this problem by simulation. Our original motivating example, cell motility in ECM with matrix remodeling under the action of matrix metalloproteinases, naturally leads to the consideration of the analogs in higher dimensions of the one-dimensional problems we have considered. Above one dimension, rigorous arguments and mean-field analyses like those we have used appear to be unavailable. We have reported some simulations in three dimensions, which show that erosion from a single

point is qualitatively similar to a random walk with a reduced diffusivity.

ACKNOWLEDGMENT

This work was supported by an Australian Research Council (ARC) Discovery Grant.

APPENDIX: APPROACH TO EQUILIBRIUM FOR A WALK IN AN INTERVAL OF RANDOM SIZE

In the following discussion the erosion probability p_s is set to zero, so the walk takes place inside an interval of fixed but random length.

1. Exact limiting mean-square displacement

Because the distribution of cluster sizes is known, it is easy to calculate the expected value of any statistical property if we know the expected value conditioned on the cluster size. This idea was applied in one-dimensional bond percolation problems associated with discrete-space, continuous-time master equations by Odagaki and Lax [26], who were interested in the frequency-dependent electrical conductivity of disordered systems. In the present context, this idea enables us to compute exactly the long-time limiting behavior of the mean-square displacement.

In a random walk on a given one-dimensional percolation cluster, the mean-square displacement will satisfy Eq. (1) exactly so long as n is small enough that the probability of the walker on this specific cluster having encountered a boundary is zero. Beyond this time, the effect of the boundary becomes more pronounced as the probability of the walker having encountered a boundary increases and eventually, roughly speaking, the mean-square displacement saturates. To calculate

$$\langle R_\infty^2 \rangle = \lim_{n \rightarrow \infty} \langle R_n^2 \rangle$$

we need to take account of the $n \rightarrow \infty$ limiting distribution of the walker's displacement from the starting position, and of the randomness within the cluster of the starting position. In the absence of boundaries the probability that the walker will be at position l after n steps, denoted as $p_n(l)$, evolves according to

$$p_{n+1}(l) = \frac{1}{2}[p_n(l+1) + p_n(l-1)]. \quad (\text{A1})$$

We seek $u(l) = \lim_{n \rightarrow \infty} p_n(l)$. Actually, unless we impose suitable boundary conditions, the sets of sites that may be occupied on odd-numbered steps is disjoint from the set that may be occupied on even-numbered steps. Given that we abort attempted steps off the cluster of allowed sites, the even-odd oscillations problem is overcome in the long-time limit so we need only solve the difference equation

$$u(l) = \frac{1}{2}[u(l+1) + u(l-1)]. \quad (\text{A2})$$

If our cluster of size m occupies the string of sites with increasing integer coordinates $\{l_1, l_2, \dots, l_m\}$, then Eq. (A2) holds for $l_1 < l < l_m$, while the boundary conditions

$$p_{n+1}(l_1) = \frac{1}{2}[p_n(l_1+1) + p_n(l_1)], \quad (\text{A3})$$

$$p_{n+1}(l_m) = \frac{1}{2}[p_n(l_m) + p_n(l_{m-1})], \quad (\text{A4})$$

imply that $u(l_1) = u(l_2)$ and $u(l_{m-1}) = u(l_m)$. The substitution $u(l) = A\alpha^l$ yields the characteristic polynomial $\alpha^2 - 2\alpha + 1 = 0$ with unique solution $\alpha = 1$, so the general solution of the difference equation (A2) is $u(l) = c_1 + c_2 l$. Either by invoking the boundary conditions, or by noting the evident left-right symmetry in the limiting distribution, we find that $u(l)$ is constant, and to preserve normalization, we have to take

$$u(l) = \frac{1}{m}. \quad (\text{A5})$$

It may be remarked that this result can be obtained in various ways. For example, it can be interpreted as the stationary distribution or invariant measure of an aperiodic finite-dimensional Markov chain [27].

The walker's starting site (s , say) is also uniformly distributed over the cluster of allowed sites and so the limiting mean-square displacement for a walk started at random on a cluster size m is

$$\langle R_\infty^2 \rangle_m = \frac{1}{m} \sum_{s=1}^m \left[\frac{1}{m} \sum_{l=1}^m (s-l)^2 \right] = \frac{m^2 - 1}{6}. \quad (\text{A6})$$

As the size of the cluster on which the walk starts is itself random, we have

$$\begin{aligned} \langle R_\infty^2 \rangle &= \sum_{m=1}^{\infty} \text{Pr}(C_p = m) \langle R_\infty^2 \rangle_m \\ &= \sum_{m=1}^{\infty} m p^{m-1} (1-p)^2 \frac{m^2 - 1}{6} = \frac{p}{(1-p)^2}. \end{aligned} \quad (\text{A7})$$

This analytic result is compared to simulation results in Fig. 12. A very good match between the long term root-mean-square displacement of a random walk obtained by simulation and the square root of Eq. (A7) is observed.

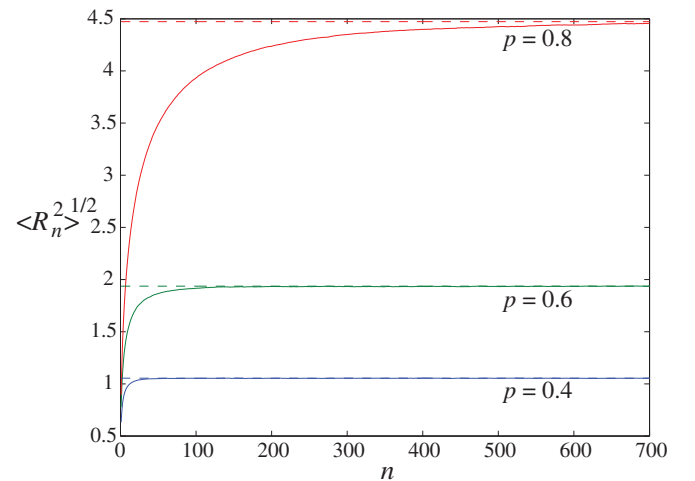


FIG. 12. (Color online) Plot of root-mean-square displacement estimated by simulations over 10^3 steps and 5×10^4 realizations for three values of p . The solid lines are simulation results and dashed lines are the square root of Eq. (A7). This shows that $\langle R_n^2 \rangle^{1/2} \rightarrow \langle R_\infty^2 \rangle^{1/2}$, given by Eq. (A7), as $n \rightarrow \infty$.

2. Simulations and empirical approximations

The calculation leading to Eq. (A7) is related to various calculations that have been done by physicists in the context of the “ant in the labyrinth” problem [4,17]. Here we wish to go further and address the transition from short-time behavior (effective ignorance of the boundaries) to long-time behavior (boundary effects strongly felt). It is evident from Fig. 12 that when $p \ll 1$, the time taken for the root-mean-square displacement $\langle R_n^2 \rangle^{1/2}$ to saturate is small, but this time becomes very long as p approaches 1.

We denote by $\langle R_n^2 \rangle_m$ the mean-square displacement after n steps, given that the cluster size C_p is m (the starting site is uniformly distributed over this cluster). Then the mean-square displacement is given by

$$\langle R_n^2 \rangle = \sum_{m=1}^{\infty} \langle R_n^2 \rangle_m \Pr\{C_p = m\}. \quad (\text{A8})$$

If the time is sufficiently short compared to the cluster size, then the walker is unlikely to have reached the boundary, so that

$$\sum_{m=1}^q mp^{m-1}(1-p)^2 \frac{m^2-1}{6} = \frac{6p - p^q[6p + \{2 + p(3-6p+p^2)\}q + 3(p-1)^2q^2 + (1-p)^3q^3]}{6(p-1)^2}, \quad (\text{A11})$$

it is straightforward to compute the approximation (A9) for any choice of m , n , and the parameter β that controls the crossover time. In Fig. 13 we compare the root-mean-square displacement obtained from simulations against estimation by the interpolation scheme defined in Eqs. (A9)–(A11). We have found empirically that taking $\beta = 2$ works best overall, and this is the value used in Fig. 13.

3. A continuum analog

To shed light on the time evolution of the mean-square displacement in our random walk on a finite interval of random length, we consider the standard one-dimensional diffusion process in a finite interval of length L with no-flux boundary conditions with the initial distribution concentrated on a single point; that is, we seek $\mathcal{P}(x, t | a, L)$ such that

$$\frac{\partial \mathcal{P}}{\partial t} = D \frac{\partial^2 \mathcal{P}}{\partial x^2}, \quad 0 < x < L, \quad (\text{A12})$$

where

$$\left. \frac{\partial \mathcal{P}}{\partial x} \right|_{x=0} = \left. \frac{\partial \mathcal{P}}{\partial x} \right|_{x=L} = 0, \quad \mathcal{P}|_{t=0} = \delta(x-a). \quad (\text{A13})$$

The standard eigenfunction expansion technique produces the solution (as a Fourier cosine series) with exponential time decay of the contributions from all of the nonconstant eigenfunctions:

$$\mathcal{P} = \frac{1}{L} + \frac{2}{L} \sum_{n=1}^{\infty} \exp\left(-\frac{n^2\pi^2 Dt}{L^2}\right) \cos\left(\frac{n\pi x}{L}\right) \cos\left(\frac{n\pi a}{L}\right).$$

Let R_t denote the displacement of the process relative to the starting position a . The mean-square displacement given the

the walk may be taken as unconstrained, leading to $\langle R_n^2 \rangle_m \approx n$. On the other hand, if the walk is sufficiently long, then the position distribution will be nearly equilibrated, so that $\langle R_n^2 \rangle_m \approx \langle R_\infty^2 \rangle_m$, where $\langle R_\infty^2 \rangle_m$ is given in Eq. (A6). For the crossover between these two regimes we choose the time n at which $m = \beta\sqrt{n}$. This leads to the approximation

$$\sqrt{\langle R_n^2 \rangle} = \sqrt{\sum_{m \leq \beta\sqrt{n}}^{\infty} \left(\frac{m^2-1}{6}\right) \Pr\{C_p = m\}} + \sqrt{n} \sum_{m > \beta\sqrt{n}}^{\infty} \Pr\{C_p = m\}. \quad (\text{A9})$$

For brevity we define $q = \lfloor \beta\sqrt{n} \rfloor$, that is, q is the largest integer that is no greater than $\beta\sqrt{n}$. Since

$$\sum_{m=q+1}^{\infty} \Pr\{C_p = m\} = p^q [1 + (1-p)q] \quad (\text{A10})$$

and

starting coordinate a and the length L (with $L \geq a$, of course) is

$$\langle R_t^2 \rangle_{a,L} = \int_0^L (x-a)^2 \mathcal{P}(x, t | a, L) dx.$$

If we now average over the starting location a (giving it a uniform distribution on $[0, L]$), then the mean-square displacement at time t in an interval of length L becomes

$$\langle R_t^2 \rangle_L = \frac{1}{L} \int_0^L \int_0^L (x-a)^2 \mathcal{P}(x, t | a, L) dx da.$$

The integrations are straightforward and we find that

$$\langle R_t^2 \rangle_L = \frac{L^2}{6} - \frac{16L^2}{\pi^4} \sum_{m=0}^{\infty} \frac{\exp[-(2m+1)^2\pi^2 Dt/L^2]}{(2m+1)^4}. \quad (\text{A14})$$

The limiting value $L^2/6$ may be compared with the analogous result (A6) for the discrete-time lattice random walk (cluster size m). We also see that the approach to the limiting distribution is very rapid.

We now make L , the length of the interval, random. The analog of the percolation-derived length distribution for the continuum problem is the Γ density,

$$\frac{d}{d\ell} \Pr\{0 \leq L \leq \ell\} = \kappa^2 \ell e^{-\kappa\ell}, \quad (\text{A15})$$

which results from taking the sum of two independent exponential random variables with mean κ^{-1} , representing the distances on the left and on the right to the nearest blocked position. Averaging Eq. (A14) using the density (A15) and making the change of variables $x = \kappa\ell$, we find that the mean-square displacement, averaged over all realizations of

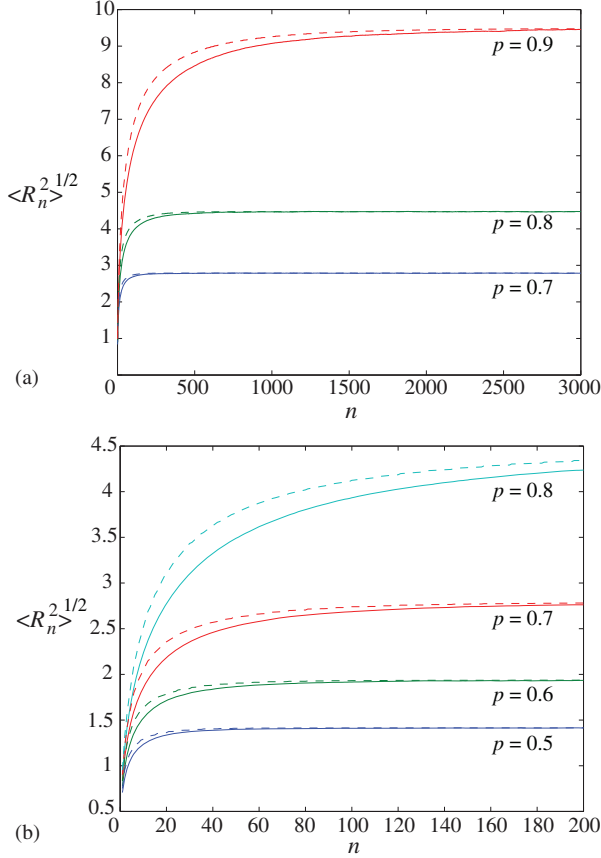


FIG. 13. (Color online) Convergence of the mean-square displacement $\langle R_n^2 \rangle^{1/2}$ to its equilibrium value, for a noneroding walker ($p_s = 0$) in a single interval of random length controlled by the allowed site density p . Solid curves show simulation data, while dashed curves show the empirical interpolation (A9) with $\beta = 2$, which we found to be the most satisfactory choice. (a) Long-time results, $0 \leq n \leq 3000$. (b) Early time results, $0 \leq n \leq 200$.

the random system, becomes

$$\begin{aligned} \langle R_t^2 \rangle &= \frac{1}{\kappa^2} - \frac{16}{\pi^4 \kappa^2} \sum_{m=0}^{\infty} \frac{1}{(2m+1)^4} \\ &\times \int_0^{\infty} x^3 \exp \left[-x - \frac{(2m+1)^2 \pi^2 \kappa^2 D t}{x^2} \right] dx. \end{aligned} \quad (\text{A16})$$

The integrals in the summand are recognized by MATHEMATICA as Meijer G functions, although this is not especially helpful. However, if we write

$$\frac{\lambda_m^3}{2} = (2m+1)^2 \pi^2 \kappa^2 D t,$$

then these integrals become

$$\begin{aligned} \mathcal{I}(\lambda_m) &= \int_0^{\infty} x^3 \exp \left[-x - \frac{\lambda_m^3}{2x^2} \right] dx \\ &= \lambda_m^4 \int_0^{\infty} z^3 \exp \left[-\lambda_m \left(z + \frac{1}{2z^2} \right) \right] dz, \end{aligned} \quad (\text{A17})$$

and so are well-suited to asymptotic estimation using Laplace's method [23]. Writing $F(z) = z + 1/(2z^2)$ we have $F'(z) = 1 - 1/z^3$ and $F''(z) = 3/z^4$, so the function $F(z)$ is concave up for $0 < z < \infty$ and attains its only minimum at $z = 1$. Hence, as $\lambda \rightarrow \infty$ we have

$$\begin{aligned} \mathcal{I}(\lambda_m) &\sim \lambda_m^4 \int_0^{\infty} z^3 \exp \left\{ -\lambda_m \left[F(1) + \frac{F''(1)}{2} (z-1)^2 \right] \right\} dz \\ &\sim \lambda_m^4 \exp \left(-\frac{3\lambda_m}{2} \right) \int_{-\infty}^{\infty} \exp \left(-\frac{3\lambda_m}{2} \zeta^2 \right) d\zeta \\ &= \lambda_m^4 \left(\frac{2\pi}{3\lambda_m} \right)^{1/2} \exp \left(-\frac{3\lambda_m}{2} \right). \end{aligned} \quad (\text{A18})$$

This leads to our identification of the long term difference between the mean-square displacement and its limiting value as a stretched exponential:

$$\begin{aligned} \langle R_t^2 \rangle &= \frac{1}{\kappa^2} \left\{ 1 - \frac{16}{\pi^4} \left(\frac{2\pi}{3} \right)^{1/2} (2\pi^2 \kappa^2 D t)^{7/6} \right. \\ &\times \left. \exp \left[-\frac{3}{2} (2\pi^2 \kappa^2 D t)^{1/3} \right] + \dots \right\}. \end{aligned} \quad (\text{A19})$$

The appearance of stretched exponentials is unsurprising, as these also arise in a mathematically closely related problem of the time to extinction for a diffusing particle released in an environment with randomly distributed traps [3,28,29].

-
- [1] K. Pearson, *Nature (London)* **72**, 294 (1905).
[2] M. N. Barber and B. W. Ninham, *Random and Restricted Walks: Theory and Applications* (Gordon and Breach, New York, London, Paris, 1970).
[3] B. D. Hughes, *Random Walks and Random Environments* (Oxford University Press, Oxford, UK, 1995), Vol. 1.
[4] B. D. Hughes, *Random Walks and Random Environments* (Oxford University Press, Oxford, UK, 1996), Vol. 2.
[5] G. H. Weiss, *Aspects and Applications of the Random Walk* (North-Holland, Amsterdam, 1994).
[6] H. Yamakawa, *Modern Theory of Polymer Solutions* (Harper and Row, New York, 1971).
[7] N. Shenvi, J. Kempe, and K. Birgitta Whaley, *Phys. Rev. A* **67**, 052307 (2003).
[8] H. Scher and E. W. Montroll, *Phys. Rev. B* **12**, 2455 (1975).
[9] K. Pearson, *Drapers' Company Research Memoirs*, Biometric Series No. III (Dulau & Co, London, 1906).
[10] A. J. Treweek, K. A. Landman, and B. D. Bell, *J. Math. Biol.* **55**, 575 (2007).
[11] A. Q. Cai, K. A. Landman, and B. D. Hughes, *Bull. Math. Biol.* **68**, 25 (2006).
[12] D. L. Nelson and M. M. Cox, *Lehninger Principles of Biochemistry* (W. H. Freeman, New York, 2010).

- [13] E. I. Deryugina and J. P. Quigley, *Cancer Metastasis Rev.* **25**, 9 (2006).
- [14] S. E. Gill and W. C. Parks, in *Extracellular Matrix Degradation*, edited by W. C. Parks and R. P. Mecham, *Biology of Extracellular Matrix No. 2* (Springer, Berlin, Heidelberg, 2011), pp. 1–22.
- [15] K. Wolf and P. Friedl, *Clin. Exp. Metastasis* **26**, 289 (2009).
- [16] M. Sahimi, *Applications of Percolation Theory* (Taylor & Francis, London, 1992).
- [17] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, 2nd ed. (Taylor & Francis, London, 1992).
- [18] M. Kac, *Ann. Math. Stat.* **16**, 62 (1945).
- [19] L. R. Shenton, *Math. Proc. Cambridge Philos. Soc.* **51**, 442 (1955).
- [20] M. Matsumoto and T. Nishimura, *ACM Trans. Model. Comput. Simul.* **8**, 3 (1998).
- [21] For a review of this topic in the context of pre-1996 work on lattice-based models, see Ref. [4]. For an account of more recent developments, see B. D. Hughes, in *Encyclopedia of Complexity and Systems Science*, edited by R. Meyers (Springer-Verlag, New York, 2009), Part 3, pp. 1395–1424. A representative collection of condensed matter physics applications of diffusion in complex environments will be found in P. Heitjans and J. Hager (eds.), *Diffusion in Condensed Matter* (Springer, Berlin, 2005). For a recent study with possible biological relevance, see D. S. Novikov, E. Fieremans, J. H. Jensen, and J. A. Helpert, *Nat. Phys.* **7**, 508 (2011).
- [22] H. Herrmann, *J. Phys. A* **16**, L611 (1983).
- [23] F. W. J. Olver, *Asymptotics and Special Functions* (Academic Press, New York, 1974).
- [24] W. Feller, *An Introduction to Probability Theory and Its Applications* (Wiley, New York, 1971), Vol. 2.
- [25] A. C. Davison and D. V. Hinkley, *Bootstrap Methods and their Application* (Cambridge University Press, Cambridge, UK, 1997).
- [26] T. Odagaki and M. Lax, *Phys. Rev. Lett.* **45**, 847 (1980).
- [27] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, 3rd ed. (Oxford University Press, Oxford, UK, 2001).
- [28] P. Grassberger and I. Procaccia, *J. Chem. Phys.* **77**, 6281 (1982).
- [29] R. F. Kayser and J. B. Hubbard, *Phys. Rev. Lett.* **51**, 79 (1983).